DOCUMENT RESUME

ED 146 780 FL 008 942

AUTHOR Andersen, Roger W.
TITLE The Impoverished State of Cross-Sectional Morpheme Acquisition/Accuracy Methodology. (or: The Leftovers Are More Nourishing Than the Main Course.) Working Papers on Bilingualism, No. 14.
INSTITUTION Ontario Inst. for Studies in Education, Toronto. Bilingual Education Project.
PUB DATE Oct 77
NOTE 39p.; For related documents, see FL 008 939-941; Paper presented at the Second Language Acquisition Research Forum (Los Angeles, California, February 13, 1977)
AVAILABLE FROM Bilingual Education Project, The Ontario Institute for Studies in Education, 252 Bloor St. West, Toronto, Ontario, Canada M5S 1V6 (as long as supply lasts).

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Bilingualism; Bilingual Students; *Data Analysis; Determiners (Languages); *English (Second Language); Form Classes (Languages); Function Words; Language Instruction; Language Research; Language Usage; *Language Variation; *Morphology (Languages); Nominals; *Puerto Ricans; Research Methodology; *Second Language Learning; Spanish; Verbs; Written Language

ABSTRACT

Current cross-sectional morpheme accuracy methodology whittles away the data until what remains in the final analysis is less interesting than what has been discarded. These methods of analysis obscure and eliminate variation in individual production of the morphemes under study and fail to reveal true systematicity in the data. Using data on eight verb-related morphemes, three "s" morphemes, and the use of articles from a study of Puerto Rican learners of English as a second language, the inadequacies in these methods are pointed out, and several improvements are suggested which will reveal whatever systematicity there is in the data without eliminating the individual variation which is so typical of learner interlanguage. (Author/CLK)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED146780

THE IMPOVERISHED STATE OF CROSS-SECTIONAL MORPHEME ACQUISITION/ACCURACY METHODOLOGY (or: THE LEFTOVERS ARE MORE NOURISHING THAN THE MAIN COURSE)

Roger W. Andersen

Colegio Universitario de Humacao
Universidad de Puerto Rico

Abstract

Current cross-sectional morpheme accuracy methodology whittles away the data until what remains in the final analysis is less interesting than what has been discarded. These methods of analysis obscure and eliminate variation in individual production of the morphemes under study and fail to reveal true systematicity in the data. Using data on eight verb-related morphemes, three "s" morphemes, and the use of articles from my study of Puerto Rican learners of English as a second language, I point out the inadequacies in these methods and suggest several improvements which will reveal whatever systematicity there is in the data without eliminating the individual variation which is so typical of learner interlanguage.

Paper presented at the Los Angeles Second Language Research Forum, UCLA. February 13, 1977. A shortened version has appeared in the Forum Proceedings, Department of English, ESL Section, UCLA.

FL008942

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Ontario Institute for Studies in Education

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

THE IMPOVERISHED STATE OF CROSS-SECTIONAL MORPHEME ACQUISITION/ACCURACY METHODOLOGY (or: THE LEFTOVERS ARE MORE NOURISHING THAN THE MAIN COURSE)

Roger W. Andersen

1. Introduction[1]

Two years ago I began looking for a way to make sense out of the great variability in performance in using English as a second language found among many Puerto Ricans who had learned English in (1) a natural English-speaking environment, (2) in English-language schools in Puerto Rico or (3) in classes in English as a Second Language (ESL) in Puerto Rico, or any combination of these. I have focussed my attention initially on those who have learned English during twelve years of compulsory ESL classes in the public schools.

In August of 1975 I obtained written samples of the English of 89 first-year university students as they were beginning their university studies. I also obtained an oral sample from 30 of these subjects, but I am using only the written data for this paper. The written data consisted of six semi-free compositions, each written on a different topic. The written instructions for writing the six short compositions (given to the subjects in Spanish) are reproduced in the Appendix, along with an English translation.

I then quantified and analyzed the data using the Group Means and the Ordering-Theoretic Methods for cross-sectional analysis, as developed by Heidi Dulay and Marina Burt (Dulay and Burt, 1974a,b). I interpret the order produced by these methods as a performance or accuracy order, as Larsen-Freeman (1975a,b) and others have pointed out. More specifically, I make no claim at this time about the universality of my findings. For the present, I consider the results of my analysis a description of the subjects' performance on this specific task. I am currently working on a comparison of this data with data obtained using different procedures. This I will report on in a future paper.

Using these methods, I soon realized that much of the apparent systematicity produced by the Group Means and Ordering-Theoretic Methods is only that - apparent - and that these methods eliminate or obscure the very variability I am attempting to study.

The thesis I will develop in this paper is that the current cross-sectional morpheme accuracy methods eliminate or obscure variability and fail to reveal true systematicity in the data. I claim that any acceptable cross-sectional method must be able to reveal true systematicity without eliminating variability. I illustrate the problem and suggest alternate ways to quantify, display and eventually explain the data.

2. Development of the thesis

In this section I discuss the performance of the 89 subjects in the use of three groups of morphemes. I follow Krashen, Madden and Bailey (1975:47-50) in using natural groups of morphemes, rather than just any group of morphemes. The three groups of morphemes are (see Table 1): (1) eight verb-related morphemes (COP, the copula be; AUX, the progressive auxiliary be; ING, the progressive -ing affix; PaI, the irregular past; PaR, the regular -ed past; Hv, the perfect auxiliary have; Vn, the past participle form in perfect aspect verb constructions; and 3RD, the present tense third-person singular s morpheme), (2) three s morphemes (PLU, the noun plural s; NsN, the possessive s morpheme; and 3RD, the present tense third-person singular s morpheme), and (3) four "articles" (A, the "indefinite" article; THE, the "definite" article; $\emptyset_1$, the absence of an article or other determiner in English in those contexts where Spanish requires a definite article; and $\emptyset_2$, the absence of an article or other determiner in English in those contexts where there is also no article or other determiner in Spanish).

Insert Table 1

2.1 Using Group Ranges to establish a group accuracy order for the morphemes

Most cross-sectional second-language (L-2) studies use the Group Means Method (Dulay and Burt, 1974a:45-6) or a similar method (for example, Fathman, 1975:35) to establish an order for the morphemes under study for the entire group of subjects. Larsen-Freeman (1975) and Rosansky (1976) and even Dulay and Burt themselves (1974b:268ff) point to the insensitivity of the Group Means to variability among the subjects. Thus, while different studies usually (although not always) report similar group orders for the morphemes, using the Group Means Method or a similar method, it is not clear what this "order" really means. This problem can be solved, in part, if we use an alternate method of determining a group accuracy order (perhaps along with the Group Means Method or a similar method for comparison) - the Group Range. The Group Range is, in effect, a modification of the 90% criterion for acquisition of a morpheme, as used by Brown (1973:258) for first-language (L-1) longitudinal research and subsequently adapted by de Villiers and de Villiers (1973) for their L-1 cross-sectional studies and, finally, by Dulay and Burt (1974a,b) for L-2 cross-sectional studies.[2]

The Group Range is illustrated in Table 2, which gives the percentage of subjects who used each of the eight verb-related morphemes correctly 90-100% of the time[3], 80-100% of the time, and 70-100% of the time, respectively (in other words, at 90% criterion, 80% criterion, and 70% criterion). (In all the calculations given in this paper, scores were counted only for those subjects who had 5 or more occasions to use the morpheme.) For any two morphemes in Table 2, the morpheme on the left

---

Insert Table 2

---

was used correctly by more subjects than the morpheme on the right, at each criterion level (with the exception of 3RD - Vn at 90% criterion).

TABLE 1

EXAMPLES OF THE MORPHEMES

(1) Verb-related	COP	The other house is cheap
	AUX	My mother is cleaning the house
	ING	My mother is cleaning the house
	PaI	I did a lot of things
	PaR	I arrived home
	Hv	I have never lived out of Puerto Rico
	Vn	I have never lived out of Puerto Rico
	3RD	He lives far from here
(2) S morphemes	PLU	two bedrooms
	NsN	my uncle's house
POSSESSIVE:	N N	my uncle house
	N of N	the house of my uncle
	3RD	He lives far from here
(3) Articles	A	It is a simple house
	THE	The other house is cheap
$\emptyset_1$(S=DEF)	LOC	Then she goes to ___ school
	TEMP	___Saturday I cleaned my house
	GEN	What do you think about ___ politics?
$\emptyset_2$(S=$\emptyset$)		He is giving ___ food to the animals

TABLE 2

ILLUSTRATION OF "GROUP RANGES". PERCENT OF SUBJECTS USING EIGHT VERB-RELATED MORPHEMES CORRECTLY AT DIFFERENT LEVELS OF ACCURACY

	COP (n=83)	AUX (n=83)	ING (n=83)	PaI (n=72)	PaR (n=40)	Hv (n=70)	3RD (n=73)	Vn (n=70)
90% CRITERION (90 - 100%)	94	53	48	31	20	11	6	6
80% CRITERION (80 - 100%)	98	66	61	45	25	14	13	7
70% CRITERION (70 - 100%)	99	70	66	53	32.5	17	16	8

Thus, this procedure gives the same group accuracy order for the eight morphemes at each of the three criterion levels. This order correlates significantly (at .01) with the order established with the Group Means Method.[4]

The superiority of the Group Range Method over the Group Means Method for establishing a group accuracy order for the morphemes is that, in addition to establishing an accuracy order for the morphemes, the Group Range tells us something about individual performance on each morpheme. (For example, 99% of the subjects used the copula correctly 70 to 100% of the time, while only 16% used third-person singular correctly at that same criterion level.) The Group Means score for a morpheme, which is the mean of all the subjects' scores for that morpheme, tells us very little about individual performance on each morpheme.

Table 2 used 90%, 80% and 70% criterion levels as a rough measure of group accuracy and variability within the group. Two alternate measures of Group Range are illustrated in Tables 3 and 4.

Insert Table 3

Table 3 gives the percentage of subjects who used each morpheme correctly 0% 1-99% and 100% of the time. The calculation of the percentage of subjects who used the morpheme correctly 100% of the time gives the same group accuracy order for the morphemes as the 90%, 80% and 70% criterion levels (with the exception of the last two, 3RD and Vn). In addition, this measure of accuracy is meaningful by itself. It shows the percentage of subjects who made no errors in using the morpheme.

Of equal interest is the observation that the calculation of the percentage of subjects who never used the morpheme correctly also gives approximately the same group accuracy order (with AUX - ING and Hv - 3RD being reversed in order). In addition, an examination of both measures (0% and 100%) along with the percent of subjects who used each morpheme correctly between 1% and 99% of the time gives a much more complete picture of the group performance on the morphemes.[5]

Table 4 shows more detail in the 1-99% range (along with the 0% and 100% levels of accuracy for comparison) by expanding the 1-99% range into intervals of 10% (except for the highest and lowest levels, which are intervals of 9%). Other variations in the calculation of the Group

---

Insert Table 4

Range are, of course, possible. The individual variation in the data is most evident in this table. For example, we can see that only 10% of the subjects omitted the copula occasionally and all of the subjects used the copula at least 60% of the time.[6] However, only half of the subjects always used the AUX when required. The rest varied from never supplying AUX (7%) to supplying it 90-99% of the time (2% of the subjects). Similar observations can be made for the rest of the morphemes in Table 4.

TABLE 3

PERCENT OF SUBJECTS USING EACH MORPHEME CORRECTLY 0%, 1-99% AND 100% OF THE TIME

	COP	AUX	ING	PaI	PaR	Hv	3RD	Vn
100%	90	51	41	24	15	11	3	6
1 - 99%	10	41	52	66	70	35	55	29
0%	0	7	6	10	15	53	42	64

TABLE 4

THE 1-99% RANGE OF CORRECT MORPHEME USE EXPANDED INTO INTERVALS OF 10% EACH

	COP	AUX	ING	PaI	PaR	Hv	3RD	Vn
100%	90	51	41	24	15	11	3	6
90-99%	4	2	7	7	5	0	3	0
80-89%	4	13	13	14	5	3	7	1
70-79%	1	4	5	8	7.5	3	3	1
60-69%	1	5	8	10	15	3	7	3
50-59%	0	7	1	1	7.5	4	7	4
40-49%	0	2	4	6	7.5	1	1	1
30-39%	0	1	2	4	5	3	7	3
20-29%	0	2	6	11	12.5	7	5	6
10-19%	0	5	5	4	5	11	10	9
1- 9%	0	0	1	1	0	0	5	1
0%	0	7	6	10	15	53	42	64

2.2 Testing the group accuracy order for individual fit

In spite of its advantages over the Group Means Method, the Group Range Method still has several limitations in revealing individual variability in performance. Using the Group Range Method we can see the percentage of subjects who fall into each range, but the method still does not show the performance of each individual in the use of the morphemes. These limitations can be overcome if we can find a procedure for testing for individual fit to the group accuracy order established with the Group Range Method. In this sense, the group accuracy order is a hypothesis that we can evaluate by seeing if the hypothesis holds for the individuals in the group. The Ordering-Theoretic Method (Dulay and Burt, 1974b:268-74; Bart and Krus, 1973) is such a procedure. It reveals a good deal of individual variability. In fact, it was adopted by Dulay and Burt for precisely this reason. Nevertheless, the Ordering-Theoretic Method has a number of serious defects which obscure the very variability it attempts to deal with.

2.2.1 The limitations of the Ordering-Theoretic Method

The Ordering-Theoretic Method (OTM) requires that all individual performance scores be converted into binary scores. This is done by using the 90% criterion: all scores between 0 and 89% become "0" (meaning "not acquired") and scores between 90 and 100% become "+" (meaning "acquired").[7] However, no justification is given for using 90% criterion instead of 80%, 70% (or even, somewhat facetiously, 83.65%) for the criterion. The 90% criterion has apparently been carried over intact from Brown (1973) and de Villiers and de Villiers (1973). (See 2.1 and footnote 2.)

The need for converting percentage scores to binary scores comes from the originators of the OTM:

> "One requirement for the prerequisite relationship to be used to interrelate items is that the items be bivalued - "1" is the item score for item i for subject k if subject k gives a correct response to item i and "0" is the item score for item i for subject k if subject k gives an incorrect response to item i." (Bart and Krus, 1973:292)

If we use this procedure of converting continuous percentage data ranging from 0% to 100% into binary "0" and "+", we are forced into claiming that a subject who uses a given morpheme correctly 80%, 86% or 89% of the time has not acquired that morpheme (or, in performance terms, does not use the morpheme correctly), while another subject who uses the same morpheme correctly 90% of the time (or 92%, 96%, etc.) has acquired it (that is, uses it correctly). This problem is illustrated in Table 5.

Insert Table 5 (see p. 56)

The OTM requires binary values for performance scores in order to test the hypothesis that morpheme X is acquired before morpheme Y. Values for morphemes X and Y, respectively, of (+ 0), (0 0) or (+ +) confirm the hypothesis and values of (0 +) disconfirm the hypothesis (Dulay and Burt 1974b:269; Bart and Krus 1973:292-3).

Table 5 shows how individual scores for two morphemes (COP and AUX) that differ by 9% for three pairs of scores in the data and by 10-11% for three additional pairs give different relationships in binary terms. For the first three pairs, this seems to be no problem at all, if we accept (+ +) and (0 0) as confirmatory relationships: the raw score for the first morpheme is larger than that for the second for all three pairs and both the raw percentage scores and the binary scores confirm the hypothesis that the first morpheme (COP) is acquired before the second (AUX)[9]. But for the second three pairs of scores, the raw scores disconfirm the hypothesis while the binary scores confirm the hypothesis in two cases ((+ +) and (0 0)) and disconfirm the hypothesis in the third case (0 +).

We must question the use of terminology appropriate to longitudinal studies ("acquired", "not acquired") when applied to cross-sectional studies, where more appropriate terms would be "correctly used", "not used correctly" and "correctly used X percent of the time". Many of the inadequacies in the OTM in current L-2 research stem from this confusion in terminology and the concommitant claims that can be made for different types of data (longitudinal acquisition data vs. cross-sectional performance data).

This problem with the use of binary scores for cross-sectional performance data becomes more evident when we consider that any sequence of scores under 90% will automatically be converted to (0 0), which in turn will be counted as a confirmation of the hypothesis; for example, using hypothetical scores, (0 89), (46 47), (16 78), (85 2), (43 0), etc. Clearly, conversion of raw percentage scores ranging from 0% to 100% into binary terms constitutes a gross distortion of the data in that it makes unfounded claims about the subjects' performance that are at odds with the subjects' actual observed performance.

If we take the OTM to be a method for determining performance hierarchies rather than acquisition hierarchies, using cross-sectional data, then we need to reconsider the acceptance of (0 0) and (+ +) as confirmatory responses and whether we are testing the right hypothesis. In performance terms, (+ 0) means that the subject performed correctly on the first morpheme and incorrectly on the second morpheme. Then (0 0) means that the subject failed to use either morpheme correctly and (+ +) means the subject used both morphemes correctly. From this perspective, we have at least two different hypotheses we can test:

(1) the score for morpheme X will be equal to or larger than the score for morpheme Y ($X \geq Y$)

(2) the score for morpheme X will be larger than the score for morpheme Y ($X > Y$) and:

(a) scores that are identical for both X and Y will not be included in the calculations

OR (b) scores that are identical for both X and Y will be included in the calculations and will be counted as disconfirmation of the hypothesis

TABLE 5

CONVERSION OF CONTINUOUS DATA (0-100%)
INTO DISCRETE TERMS (BINARY SCORES)

PERCENTAGE POINTS DIFFERENCE	RAW SCORE COP	RAW SCORE AUX	BINARY SCORE COP	BINARY SCORE AUX
COP LARGER THAN AUX BY:				
9%	100	91	+	+
9%	96	87	+	0
9%	89	80	0	0
AUX LARGER THAN COP BY:				
10%	90	100	+	+
11%	89	100	0	+
11%	77	88	0	0

COP, AUX = "PERFECT" COP and AUX
90-100 = + 0-89 = 0

TABLE 6

TESTING HYPOTHESES WITH BINARY SCORES AND RAW PERCENTAGE SCORES

HYPOTHESIS:			COP ≥ AUX			COP > AUX		
			CONF.	DISC.		CONF.	DISC.	PERCENT COP=AUX
6A								
AUX	BINARY		97%	3%	(N=34)	94%	6%	56%
COP	SCORES	(N=78)						++ OR 00
OR	RAW		97%	3%	(N=39)	95%	5%	50%
AUX	SCORES							100-100
6B								
PERFECT	BINARY		94%	6%	(N=39)	87%	13%	50%
COP	SCORES	(N=78)						++ OR 00
OR	RAW		85%	15%	(N=58)	79%	21%	26%
AUX	SCORES							100-100

For the first hypothesis, (+ 0), (+ +) and (0 0) are confirmatory responses and (0 +) disconfirms the hypothesis. For hypotheses 2a and 2b, (+ 0) confirms the hypothesis and (0 +) disconfirms it. (+ +) and (0 0) are not included in the calculation for hypothesis 2a, whereas for hypothesis 2b they are counted as disconfirmatory responses. The OTM tests only the first hypothesis.[10] In the rest of this paper I will use version 2a of hypothesis 2 whenever I use raw percentage scores.

Table 6 illustrates the testing of hypotheses 1 and 2a with binary scores and raw percentage scores for the morphemes COP and AUX. Hypothesis 1

Insert Table 6

---

is given as COP $\geq$ AUX and hypothesis 2a as COP $>$ AUX. The table gives the percent of confirmation and the percent of disconfirmation using binary scores and raw scores for both hypotheses, as well as the percentage of subjects who had identical scores for both morphemes, both (+ +) or (0 0) for binary scores or both 100% for raw percentage scores.[11]

6A illustrates this for ANY COP and AUX, meaning that an occurrence of COP or AUX is scored as correct if it is supplied in an obligatory context, even if the form of the COP or AUX is marked incorrectly for tense or person-number agreement. For example, *he are* for *he is*; *they was* for *they were* or *they are* are scored as "used morpheme correctly". Scoring this way credits the subjects for knowing that a copula or auxiliary is required, but not necessarily for knowing rules of tense and person-number agreement. 6B gives the percent of confirmation and disconfirmation for PERFECT COP and AUX, meaning that only the perfectly correct form of COP or AUX is acceptable. Here, in addition to knowing that the morpheme is required, the subject must know the tense and person-number agreement rules.

In 6A we see that for ANY COP and AUX we get basically the same results using binary scores and using raw percentage scores. But in 6B we get a much higher percent of disconfirmation with raw percentage scores than with binary scores (15% vs 6% for COP $\geq$ AUX; 21% vs 13% for COP $>$ AUX). Approximately half of the subjects had identical binary scores for both morphemes in 6A and 6B, but fewer subjects in 6A (only 26%) had identical raw percentage scores than in 6B (where 50% of the subjects had identical scores). This suggests that the different results for binary scores and raw scores in 6B (in contrast with the similar results for both methods in 6A) are due to the way in which the binary scoring procedure converts different pairs of scores (in percentage) into identical pairs of binary scores ((+ +) or (0 0)). This is clear evidence that the use of binary scores distorts the data.

If we are concerned with comparing the subjects' use of ANY copula or auxiliary with their use of PERFECT copula and auxiliary, we can see that the use of binary scores glosses over the very real difference that exists between the two. For the hypothesis COP $\geq$ AUX, this is a difference (between ANY and PERFECT use of the morphemes) of only 3 percentage points using binary scores but 12 percentage points using raw scores. For the hypothesis COP $>$ AUX, we find a difference of 7 percentage points using binary scores, but 16 percentage points using raw scores. Since the raw scores represent the actual performance of the individuals, we can only conclude that, at least for this set of data, binary scoring fails to reveal accurately the subjects' ability to use COP and AUX with correct

tense and person-number agreement (as opposed to their ability to use ANY form of COP and AUX).[12]

Dulay and Burt (1974a:43-4) use a scoring procedure which gives half-credit for certain misformed morphemes. The preceding discussion shows that we do not need to weight scores in this way if we can separate out the factors that would lead to such weighting. In the case of COP and AUX the use of ANY form of the morphemes was calculated separately from the use of tense marking and subject-verb agreement for person and number (PERFECT AUX and COP – henceforth abbreviated as PF AUX and PF COP). I believe that all the cases of weighted scores in Dulay and Burt (1974a) can be handled this way and recommend that weighted scores not be included in the analysis of the data without ample justification. Instead, the factors leading to this weighting should be studied in more detail and scored separately.

2.2.2 *Testing individual fit with raw percentage scores*

Table 7A shows the group accuracy order (hypothesis 2a) for the three morphemes COP, AUX and ING (both ANY COP and AUX and PF COP and AUX) as established by the Group Range Method. The four criterion levels

---

Insert Table 7

(100%, 90%, 80% and 70%) give the same orders: ANY COP > ANY AUX > ING and PF COP > ING > PF AUX. These scores indicate that, when tense marking and person-number agreement are disregarded, more subjects performed well (at each criterion level) using COP than using AUX and more performed well using both COP and AUX than using ING, but that when we take tense-marking and person-number agreement into consideration, the order of AUX and ING reverses. We could hypothesize, based on these figures, that the additional burden of using the forms correctly marked for tense and person-number agreement makes it harder to learn and/or use the AUX than to learn and/or use ING correctly, thus accounting for ING preceding AUX.[13]

These two orders constitute hypotheses about the performance orders of individuals which can be tested for individual fit by examining the raw percentage scores for each of the individuals. For example, for an individual who used ANY COP 100% of the time, ANY AUX 91% of the time and ING 64% of the time, the hypothesis holds. (This same individual had PF COP 96%, PF AUX 73%, ING 64%, which does not fit the group hypothesis PF COP > ING > PF AUX.)

Table 7B shows the results of the test for individual fit to the group hypothesis (2a). The percent of individuals who conform to the group order is given first for the order between any two morphemes and then for all three morphemes together. Finally, the percent of individuals who conform to the hypothesis X $\geq$ Y $\geq$ Z (hypothesis 1) for the three morphemes is given for comparison. This is done for both ANY COP and AUX (on the left) and PF COP and AUX (on the right).

As we saw in Table 6, a larger percentage of individuals conform to the performance order for ANY COP and AUX and ING than for PF COP and AUX and ING. This greater individual variation in the use of these morphemes is probably due to the increased complexity of the task: when the subject

TABLE 7

7A: THE GROUP ACCURACY ORDER FOR COP, AUX and ING

	ANY COP >	ANY AUX >	ING	PF COP >	ING >	PF AUX
100%	90	51	41	55	41	36
90-100%	94	53	48	77	48	38
80-100%	98	66	61	90	61	55
70-100%	99	70	66	92	66	65

7B: TEST FOR INDIVIDUAL FIT - PERCENT OF INDIVIDUALS WHO CONFORM TO THE GROUP ACCURACY ORDER

ANY COP >	ANY AUX >	ING			PF COP >	ING >	PF AUX	
ANY COP >	ANY AUX		95%	(37/39)	PF COP >		PF AUX	79% (46/58)
ANY COP >		ING	92%	(45/49)	PF COP >	ING		72% (41/57)
	ANY AUX >	ING	55%	(26/47)		ING >	PF AUX	62% (34/55)
ANY COP >	ANY AUX >	ING	28%	(14/50)	PF COP >	ING >	PF AUX	25% (15/61)
ANY COP ≥	ANY AUX ≥	ING	77%	(60/78)	PF COP ≥	ING ≥	PF AUX	58% (45/78)

must know not only when to use a COP or AUX but also exactly which form to use, we would expect less accurate performance (and learning) and greater individual variation in using one morpheme more accurately than the other. (It is this type of variability that is obscured by using binary scoring instead of the raw percentage scores we have used in this test.)

It was mentioned in 2.2.1 that the OTM tests only hypothesis 1: $X \geq Y$. In addition, the OTM tests only the relationship between any two items. Yet the "tree" made from these relationships gives the impression that the OTM tests the individual fit to a hypothesis involving up to four items, such as $Q \geq R \geq S \geq T$ (Bart and Krus, 1973:207; Dulay and Burt, 1974b:270-73). This is extremely misleading, as Table 7B reveals. We cannot assume that a confirmed relationship between any two of three morphemes, X, Y and Z, such that $X > Y$, $X > Z$ and $Y > Z$ automatically implies that $X > Y > Z$. If we replace X, Y and Z with ANY COP, ANY AUX and ING, respectively, we can see in Table 7B that 95% of the subjects fit the hypothesis ANY COP > ANY AUX, 92% fit the hypothesis ANY COP > ING, only 55% fit the hypothesis ANY AUX > ING and only 28% of the subjects fit the hypothesis ANY COP > ANY AUX > ING. Thus the high degree of fit of the individual subjects to the first two hypotheses in no way implies that the hypothesis ANY COP > ANY AUX > ING will hold. Although for some other group of three morphemes the degree of fit for $X > Y$, $X > Z$ and $Y > Z$ may be approximately the same as for $X > Y > Z$, this cannot be assumed, but must be tested.

From Table 7B, then, we can see that the degree of individual fit to the group hypotheses ANY COP > ANY AUX and ANY AUX > ING is close to perfect (95% and 92%).[14] The low percentage of individuals who conform to the hypothesis ANY AUX > ING (55%) and ING > PF AUX (62%), as well as the even lower percentages for ANY COP > ANY AUX > ING (28%) and PF COP > ING > PF AUX (25%) suggests that we should examine the data for the individual subjects for an explanation.

2.2.3 A closer look at the data for COP, AUX and ING

Up to this point we have followed other second-language researchers (who in turn based their work on Brown 1973) in assuming that AUX and ING are two independently-learned morphemes. It is quite possible that they are learned independently of each other in beginning stages, but for most of the subjects in this study the two morphemes are clearly interrelated. Table 8 gives the percent of subjects who used each of four different verb forms in a progressive context: AUX-ING[15] (he is working), ING (he working), AUX-V (he is work) and V (he work or, occasionally, he works or he worked).

Insert Table 8

From Table 8 we can see that the target for most of the subjects is the complete form AUX-ING, rather than ING, AUX or V alone. In fact, 35% of the subjects used AUX-ING 100% of the time and 60% used it 70-100% of the time. This explains, in part, the difficulty we encountered in Table 7B in confirming the two group hypotheses ANY COP > ANY AUX > ING and PF COP > ING > PF AUX: since a large percentage of the subjects are using both AUX and ING together, it is obviously difficult, if not impossible, to establish an order for them.

TABLE 8

VERB FORMS USED IN PROGRESSIVE CONTEXT

PERCENT OF SUBJECTS USING EACH FORM

	AUX-ING	ING	AUX-V	V
100%	35	1	1	1
1 - 99%	50	26	35	39
0%	14	72	64	60
70 -100%	60	1	3	8

This is even more evident in Figure 1, which gives the performance of each subject in the use of AUX-ING, ING, AUX-V and V in a progressive context. The subjects' scores are grouped into 4 vertical columns (1-4)

Insert Figure 1

---

and 4 horizontal divisions (A-D). The subjects in divisions A and B used AUX-ING at least part of the time (from 14% to 100% of the time). Those in A never used V as an alternative, while those in B sometimes did (11% to 83% of the time) along with the other forms. The much smaller groups in C and D never used AUX-ING, but used one or more of the other forms (ING, AUX-V or V). D shows that only one subject consistently marked the progressive with AUX alone; only one consistently used ING by itself; and only one never used any progressive marker. Nine subjects are grouped under C, those who never used AUX-ING, but used at least two of the other alternative forms.

Column 1 gives those subjects who never used ING or AUX alone; column 2, those who used AUX alone, but never ING without AUX; column 3, those who used ING alone, but never used AUX without ING; and column 4, those who used both ING alone and AUX alone. The number of subjects in each cell is as follows:

		COLUMNS				
		1	2	3	4	TOTAL
GROUPS	A	29	11	5	1	46
	B	8	5	6	3	22
	C	0	3	3	3	9
	D	0	1	1	1	3
	TOTAL	37	20	15	8	80

This matrix, together with Table 8 and Figure 1, shows that the ANY AUX > ING order shown in Table 7A is primarily due to the fact that there are more individuals who use AUX without ING to express progressive than individuals who use ING without AUX. (Notice also in column 4 that of seven subjects who used ING by itself and AUX by itself part of the time, five used AUX by itself more frequently.) Although the numbers we are dealing with are small, it seems that the free morpheme AUX is "picked up" and used as a progressive marker more often than the bound morpheme ING.[16]

The individual variability seen in Table 7B is even more evident in Figure 1, where we can see that this variability is systematic.[17] As part of this picture of variability and systematicity, we can see the following: (1) although performance on AUX-ING varies from 0% to 100%, most of the subjects (68 of the 80 subjects in Fig. 1) use AUX-ING at least part of the time (groups A and B). (2) These 68 subjects fall into two smaller groups: A, those who never use an unmarked verb as an alternative and B, those who sometimes use an unmarked verb. (It seems possible that

	AUX ING	1 ING	AUX	V	AUX ING	2 ING	AUX	V	AUX ING	3 ING	AUX	V	AUX ING	4 ING	AUX	V
A:	100	0	0	0	94	0	6	0	89	11	0	0	56	11	33	0
	100	0	0	0	91	0	9	0	88	12	0	0				
	100	0	0	0	90	0	10	0	80	20	0	0				
	100	0	0	0	88	0	12	0	80	20	0	0				
	100	0	0	0	88	0	12	0	57	43	0	0				
	100	0	0	0	88	0	12	0								
	100	0	0	0	86	0	14	0								
	100	0	0	0	83	0	17	0								
	100	0	0	0	80	0	20	0								
	100	0	0	0	80	0	20	0								
	100	0	0	0	67	0	33	0								
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
	100	0	0	0												
B:	89	0	0	11	80	0	7	13	84	8	0	8	55	18	9	18
	88	0	0	12	64	0	27	9	57	14	0	29	33	33	17	17
	88	0	0	12	45	0	45	9	56	22	0	22	17	8	33	42
	75	0	0	25	43	0	43	14	33	11	0	55				
	70	0	0	30	14	0	43	43	20	60	0	20				
	64	0	0	36					14	14	0	71				
	40	0	0	60												
	17	0	0	83												
C:					0	0	88	12	0	20	0	80	0	25	75	0
					0	0	63	37	0	13	0	87	0	20	40	40
					0	0	13	87	0	8	0	92	0	18	27	55
D:					0	0	100	0	0	100	0	0	0	0	0	100

FIGURE 1 VERB FORMS USED IN PROGRESSIVE CONTEXT BY 80 SUBJECTS

Read horizontally, from left to right: for example, the first subject in column 2, group A used AUX-ING 94% of the time, ING alone 0%, AUX alone 6%, and neither AUX nor ING (V) 0% of the time.

(Number of subjects in each cell.)

	1	2	3	4	TOTAL
A	29	11	5	1	46
B	8	5	6	3	22
C	0	3	3	3	9
D	0	1	1	1	3
TOTAL:	37	20	15	8	80

those in B "know" the correct progressive form, but have not yet learned all the contexts which require it.) (3) Although it is logically possible that many subjects would consistently use AUX without ING (cell D-2) or ING without AUX (cell D-3) or the unmarked verb (cell D-4) in a progressive context, only one subject in each category actually performed this way.[18] This type of display of data, then, especially in conjunction with devices such as those used in Tables 7 and 8, is capable of revealing true systematicity and variability at the same time.

2.2.4 Variability and systematicity in the possessive NsN construction

We can illustrate the procedures used for analyzing the progressive construction by examining the possessive noun + 's + noun construction (NsN) in comparison with two other s morphemes - PLU and 3RD. (See Table 1 for examples of these morphemes as well as of two other constructions often used by the subjects to express possession: noun + noun (NN) and noun + of + noun (N of N).[19] Table 9A gives the group accuracy order using the Group Range Method for the three morphemes, PLU > NsN > 3RD, at each of five different criterion levels. The results of the test for individual fit to the group hypothesis are given in 9B. Although the

---

Insert Table 9

---

percentage of individuals who conform to the group order for any two of the morphemes is high,[20] the individual fit to the order for all three morphemes (49%) is not at all impressive. Even for hypothesis 1 of PLU $\geq$ NsN $\geq$ 3RD, the degree of fit is still low.

These results lead us into a consideration of the variability in the data. While PLU and 3RD are simple bound morphemes, NsN is not. This is evident in the examples in Table 1 of the forms used to express possessive: NsN, NN and N of N. The so-called possessive morpheme 's is added to the syntactic NN construction. The percent of subjects who used this NN construction (with or without the possessive 's) is shown in Table 9A as $\frac{\text{NsN}}{\text{NN}}$. If we take the order of the two nouns as the basic indicator of possession, then possessive precedes both PLU and 3RD. If we assume that the subject will use the NN construction and only count the percentage of times he supplied 's in the NN construction, we get the order PLU > NsN > 3RD. We can see, however, that the possessive construction is syntactically more complex than either PLU or 3RD and as such is more subject to individual variability.

This variation is more evident in Table 10, which gives the percent of subjects who used each of the three forms, NsN, NN and N of N, to express possession (as well as the percentage of individuals who used NN or NsN - i.e., who used a noun + noun construction, with or without the 's). As we found with the progressive, the native construction (for

---

Insert Table 10

---

progressive, AUX-ING; for possessive, NsN) is the target construction more often than any of the other non-native alternatives. But the degree of individual variation is still great. This is probably due to both inter- and intralingual overgeneralization. The Spanish possessive noun + de + noun (possessed + de + possessor) parallels the English

TABLE 9

9A: THE GROUP ACCURACY ORDER FOR PLU, NsN and 3RD

	PLU >	NsN >	3RD	NsN N N
100%	20	15	3	41
90-100%	35	21	6	47
80-100%	55	31	13	58
70-100%	65	35	16	67
0%	2	21	41	4

9B: TEST FOR INDIVIDUAL FIT - PERCENT OF INDIVIDUALS WHO CONFORM TO THE GROUP ACCURACY ORDER

PLU	>	NsN	> 3RD	
PLU	>		3RD	99%
PLU	>	NsN		81%
		NsN	> 3RD	71%
PLU	>	NsN	> 3RD	49%
PLU	≥	NsN	≥ 3RD	60%

TABLE 10

FORMS USED TO EXPRESS "POSSESSIVE"

PERCENT OF SUBJECTS USING EACH FORM

	NsN or N N	NsN	N N	N of N
100%	41	15	6	4
1-99%	55	64	52	56
0%	4	21	42	40
70-100%	67	35	16	14

noun + of + noun construction, which is sometimes used for possession, but also for other situations. The subject may be following a Spanish model in some cases and an English model in others. For example, John's friend and a friend of John's would both be N de N constructions in Spanish: el amigo de Juan, un amigo de Juan. This same N de N construction in Spanish is the equivalent of many noun adjunct constructions in English. For example, consider una casa de ladrillos - a brick house vs. un montón de ladrillos - a pile of bricks. Thus there are several sources within English and in Spanish for both the NN and N of N constructions.

Figure 2 presents the scores for the three possessive forms for the 52 subjects who had 5 or more obligatory occasions to express possession. They are grouped into three categories, so as to show the degree of variability and systematicity in the data at the same time. 21 subjects (on the left) never used N of N but consistently used the NN construction (with or without 's) to express possession. Of these, 8 used only the complete NsN form and 3 used only the NN form. The remaining 10 subjects in this group used both NsN and NN to varying degrees. Thus, in spite of the obvious variability, all 21 subjects are consistently using a NN construction to express possession.

Insert Figure 2

The second group of 18 subjects used the N of N construction (from 6% to 82% of the time) along with either the NsN (12 subjects) or the NN construction (6 subjects). The 12 subjects who used both NsN and N of N to express possession seem to have "learned" the NsN possessive form but have varying degrees of success at drawing the line between when to use NsN and when to use N of N. As mentioned previously, they could be misled by Spanish or, more likely, the many N of N constructions in English (the bottom of the page, for example). The smaller group of 6 subjects who used NN and N of N to express possession seem to have taken NN as the marker of possession and, like the previous group of 12 subjects, have difficulty in deciding when NN is required instead of N of N.

Two very different types of subjects were grouped together in the last grouping, partly for lack of space: 11 subjects who used all three constructions to express possession (arranged in increasing use of N of N) and two subjects who only used N of N.[21] The last two subjects, it seems, have taken the construction most similar to Spanish to express possession. The other 11 subjects use all three forms to some degree. If we take a closer look at certain interlingual correspondences between Spanish and English, we see a possible reason for this variable use of all three forms. There are numerous cases where Spanish N de N is equivalent to English NsN, NN or N of N. For example:

SPANISH			ENGLISH							
N	DE	N		N 's	N	N	N		N OF	N
el amigo	de	Juan	=	Juan's	friend					
una casa	de	ladrillos	=			a brick	house			
una lista	de	libros	=			a book	list /	a	list of	books
el sótano	de	la casa	=					the	basement of	the house

21			18						13		
			12			6					
NsN	N N	N of N	NsN	N N	N of N	NsN	N N	N of N	NsN	N N	N of N
100	0	0	94	0	6						
100	0	0							60	33	7
100	0	0	90	0	10						
100	0	0							56	33	11
100	0	0	88	0	12						
100	0	0							43	43	14
100	0	0							23	64	14
100	0	0	80	0	20	0	80	20			
									62	15	23
						0	75	25	13	62	25
									29	42	29
92	8	0	70	0	30						
88	12	0	67	0	33						
80	20	0	60	0	40				20	40	40
80	20	0	55	0	45						
75	25	0				0	53	47			
44	55	0				0	50	50	17	33	50
40	60	0	44	0	56						
33	67	0				0	43	57			
30	70	0	40	0	60						
20	80	0							9	18	73
			25	0	75						
						0	20	80			
			18	0	82				6	12	82
0	100	0									
0	100	0							0	0	100
0	100	0							0	0	100

FIGURE 2 FORMS USED TO EXPRESS "POSSESSIVE" FOR THE 52 SUBJECTS WITH N $\geq$ 5

The Spanish surface form is no clue to which English construction is required.

2.2.5 Discussion

I have shown that the Group Range is superior to the Group Means for establishing a group accuracy order because the Group Range is, at the same time, a partial measure of individual variability in the data. The test for individual fit to the group order established with the Group Range went further in revealing this variability. Finally, the graphic display of individual performance scores in Figures 1 and 2 went still further in showing simultaneously whatever systematicity and variability was evident in the subjects' performance. If we have only the Group Means Method and the Ordering-Theoretic Method at our disposal, we are severely limited in what we can learn from the data.

2.3 How to know which morphemes to study - let the data guide the researcher

Brown (1973:264) found it necessary to combine the articles *a* and *the* into a single morpheme, ART(icle). (See "articles" in Table 1 for examples.) The de Villiers (1973) and Dulay and Burt (1974a,b) carried this decision over into their work, evidently because they were trying to replicate Brown's results. Nevertheless, there is good reason to question this decision to lump together two related but still quite distinct morphemes.[22] In fact, the data used for this study suggests strongly that *the* and *a* are very different learning and/or performance problems.

Figure 3 shows the percentage of each morpheme (*a* and *the*) supplied in obligatory context (plotted on the ordinate) for each subject (plotted on the abscissa). It is clear from this figure that *the* and *a* constitute

---

Insert Figure 3

---

very different learning and/or performance problems for these subjects. Most of the subjects perform quite well with *the* in obligatory contexts, but their performance with *a* ranges from almost 0% up to 100%. To combine *a* with *the* as ART would definitely obscure the subjects' use of these morphemes.

It is interesting to note that the English definite article *the* is obligatory in more or less the same contexts that require the definite article in Spanish.[23] With *a*, however, the picture is not so simple. A more detailed analysis than space permits here would reveal that many of the subjects may be following a strategy of using the English equivalent of whatever article is required in Spanish in that context. For the definite article, few errors result from this strategy. For the indefinite article, however, the strategy will produce a significant number of errors. Spanish often requires no article where English requires *a*. In addition, Spanish *un, una* corresponds to both *a* and *one* in English. Many of the errors in the use of *a* seem to be due to this strategy: the article *a* is frequently omitted in contexts where it is omitted in Spanish and *one* is often used where *a* is required.

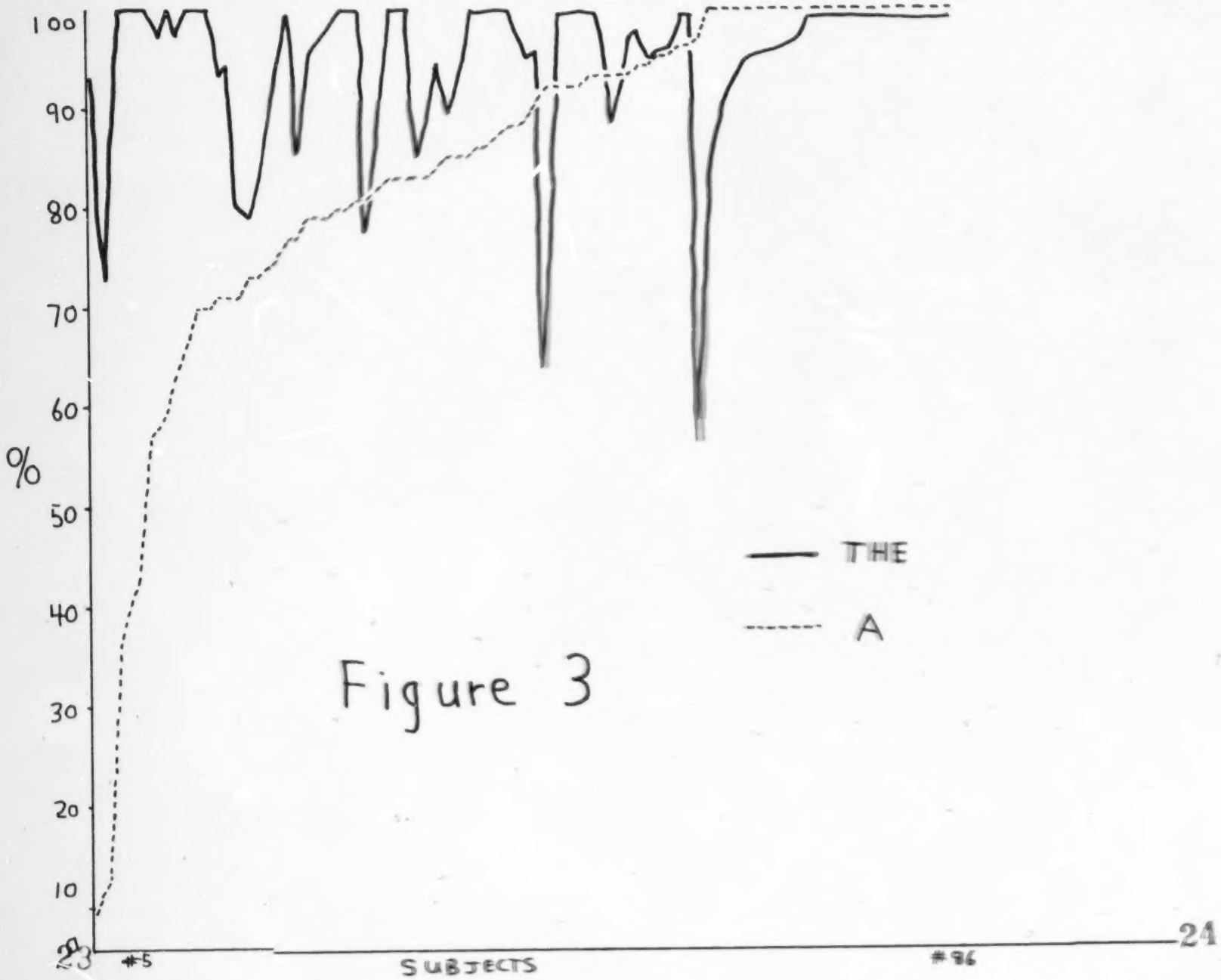

Figure 3

Other errors in the use of the articles result from overcompensation, a type of overgeneralization. A subject will systematically leave out an otherwise correct *the* in contexts where Spanish and English both require the definite article, seemingly from an awareness that English often omits the definite article in contexts where Spanish requires it. If *the* and *a* were combined as ART, we would not be able to make these important observations.[24]

A possibly useful "principle" of L-2 research that would keep the researcher from overlooking these important details would be something like the following: let the data determine which morphemes - or, more generally, which aspects of learner interlanguage - we should study. This "principle" is merely a reminder to the researcher. As Schachter (1974) has shown, not everything a researcher wants to know about the subjects' use of grammatical constructions is evident in performance data.

2.4 The notion of "obligatory context" for a morpheme

The notion of "obligatory context" for the use of a morpheme, an extremely important one for both longitudinal and cross-sectional morpheme acquisition/accuracy studies, originated with Brown (1973:255) for current L-2 research. Hakuta (1974, 1975) has shown (with longitudinal data) that there is more to the use of a morpheme than correct use in a context which requires that morpheme. The use of other morphemes in that context and the use of the morpheme in the wrong context are also important in finding out how learners learn and use their second language. Brown himself went far beyond mere "correct use of a morpheme in obligatory context" in his L-1 research. The point I wish to make here is that cross-sectional L-2 morpheme research has not gone much beyond the original notion of "obligatory context".

The data from the subjects in this study provides strong evidence for the need to go beyond the notion of correct use of a morpheme in obligatory contexts. One aspect of this evidence is the use of *the* in obligatory contexts as well as in contexts where *the* is not permitted. As was seen in Figure 3, most of the subjects perform very well in using *the* when it is required. But what is not at all evident if we only consider "obligatory contexts" is that they also use *the* when it is *not* required. Figure 4 shows the percentage of *the* supplied in obligatory contexts (plotted on the ordinate) for each subject (plotted on the abscissa) along with the percentage of $\emptyset_1$ (= no article) "supplied" in contexts where English does not permit but Spanish requires a definite article. (See "articles" in Table 1 for examples of $\emptyset_1$, divided into three general groups: locative exceptions to the use of the definite article, temporal exceptions, and generic reference.)

Insert Figure 4

The subjects vary considerably (from 0% to 92% correct) in their use of "zero article" ($\emptyset_1$) in contexts that require it in English but require the definite article in Spanish. Almost all of the errors made in these cases consist of using *the* instead of no article. As in the case of *a*, it seems that the subjects are following a strategy of using

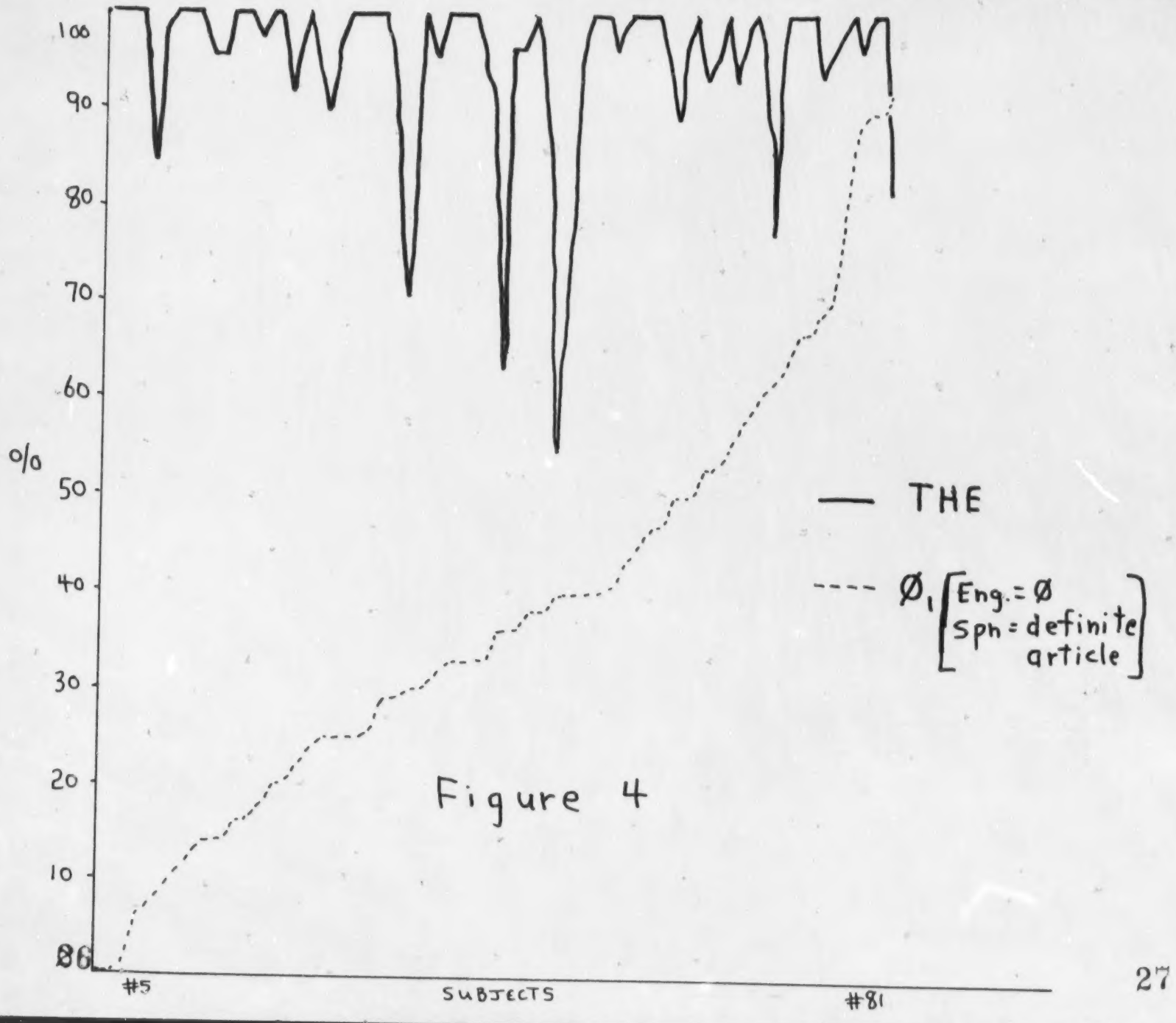
100
90
80
70
60
%
50
40
30
20
10
Ø
#5
SUBJECTS
#81
THE
Ø1 [Eng. = Ø
Spn = definite article]

Figure 4

the English equivalent of whatever article is required in Spanish in that context. That this is probably the case is seen in Figure 5, which compares the same percentage of $\emptyset_1$ "supplied" in obligatory contexts with $\emptyset_2$, which is "zero article" in those contexts where neither English nor Spanish requires an article. Although there are several gaps in the data (several subjects had fewer than 5 obligatory occasions for $\emptyset_2$), it is clear that the subjects perform with $\emptyset_2$ as well as with obligatory the. Thus, it seems that wherever English and Spanish both require the definite article or do not permit an article (the and $\emptyset_2$) the subjects perform well. When Spanish and English require different articles (a and $\emptyset_1$) considerable variability is evident and many of the errors seem to be interference errors.[25]

Insert Figure 5

This difference between "correct use of a morpheme in obligatory contexts" and real control over the use of the morpheme is evident only if we refine the notion of "obligatory context" to include the incorrect use of the morpheme in contexts where it is not permitted and a consideration of which morpheme is used when the required morpheme is not. Thus, for a subject who has apparently learned a given morpheme, we need to study how well he has learned the restrictions on the use of that morpheme as well.

3 Summary and conclusions

The techniques I have illustrated in this paper may not be precisely the best possible techniques for cross-sectional L-2 methodology. I have used them here with two purposes: (1) to suggest them as potential improvements in current cross-sectional L-2 methodology and (2) to emphasize the need for analytical techniques that will not misrepresent the raw data and which will reveal whatever systematicity and variability is present in the data. The several points I have tried to make can be summarized as follows:

(1) Group Ranges can be used (perhaps along with Group Means and other methods) to establish a group accuracy order for the morphemes being studied. The advantage of Ranges is that they permit a display of finer detail of variability while revealing, at the same time, the same group accuracy order given by the Group Means (at least for this data).

(2) No justification has been given for cross-sectional L-2 research for accepting 90% (or 80%, 70%, etc.) as a cut-off point for "correct use" of a morpheme or for converting performance scores to binary terms. If these decisions to reduce continuous data to discrete terms are to be made, they must be justified. If they cannot be justified, we must look for other procedures for quantifying the data.

(3) One such procedure is the use of raw percentage scores to test individual fit to group accuracy orders given by some method like the Group Ranges Method.

(4) The Ordering-Theoretic Method is really testing the hypothesis "the score for morpheme X is as large as or larger than the score for morpheme Y". We should also test the hypothesis

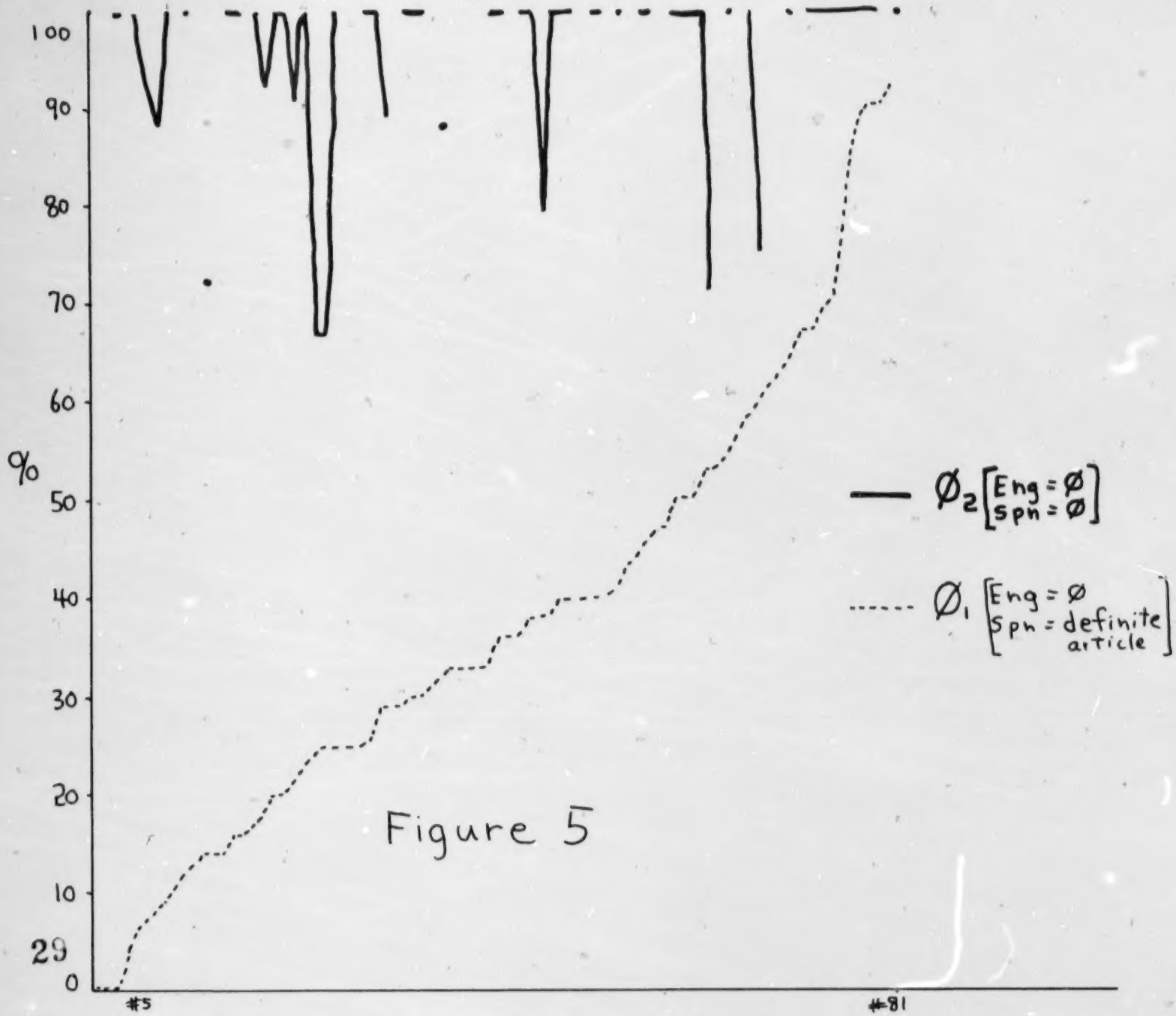
100
90
80
70
60
%
50
40
30
20
10
0
#5
#81
$\emptyset_2$ [Eng = ∅, Spn = ∅]
$\emptyset_1$ [Eng = ∅, Spn = definite article]

Figure 5

that the individual score for morpheme X is larger than that for morpheme Y.

(5) The ordering-Theoretic Method tests only the individual fit to a hypothesis that any two morphemes are ordered ($X \geq Y$, $X \geq Z$, $Y \geq Z$, etc.). It appears to be testing hypotheses such as $X \geq Y \geq Z$, but it really isn't, since the individual fit to this hypothesis is never checked independently. We need to use a procedure that will check individual fit to the hypotheses $X \geq Y \geq Z$, $X > Y > Z$ and others like these, as well as hypotheses concerning the ordering of any two morphemes.

(6) We need to study the accuracy order of not just any group of morphemes, but rather of natural groups of morphemes, such as NP-related morphemes as one group, V-related morphemes as another (Krashen, Madden and Bailey 1975).

(7) When possible we should separate out factors that would lead to weighting scores. They may point to aspects of learning and/or using English that we had overlooked.

(8) The degree of individual fit, expressed as the percent of subjects who have morpheme accuracy scores which fit the group hypothesis, while perhaps not statistically significant for a given test, could nevertheless reveal important patterning in the data.

(9) The results of the test of individual fit may lead to better ways of displaying the data so as to reveal systematicity and variability simultaneously.

(10) It is not always obvious which morphemes - or, more generally, which aspects of learner language - we should study. We should not allow the precedent established in previous L-1 and L-2 research to bias our decisions. Instead we should let the data itself guide us.

(11) The notion of "obligatory context" for use of a morpheme needs refining.

These points by no means exhaust the possibilities of alternate techniques for analyzing, quantifying and displaying cross-sectional data. Researchers dealing with cross-sectional data would do well to draw on the techniques used in longitudinal studies and in other fields, especially other areas of psycholinguistics and in sociolinguistics.

Basically I have suggested that cross-sectional methodology be more accountable to the data, but in retrievable form that will allow us wide ranges of alternate representations of the data. It is often necessary to abstract away from the data and, in doing so, to convert continuous performance data into discrete terms. This conversion should not be as mechanical and as arbitrary as it has been. Nor should it always be necessary. In addition, we need to go beyond the study of morphemes in isolation and relate the knowledge of these morphemes to

the knowledge of constructions that employ the morphemes and we need to get a broader view of a learner's use of a wide number of grammatical constructions. Once we have this information we can attempt to relate it to non-linguistic facts about the learner , such as his attitude towards the language, himself and speakers of the language; years in school; contact with native speakers of the language; performance in his first language; fluency in the second language, etc.

FOOTNOTES

1. I am indebted to Heidi Dulay and Marina Burt, whose pioneer work in cross-sectional second-language morpheme acquisition research formed the basis for all subsequent studies, including this one, and to numerous other second-language researchers who, either directly or indirectly, contributed to my perception of the nature of this research and the methodology needed for it.

2. The choice of the 90% criterion for L-1 longitudinal research was fully justified by Brown (1973:258) and we can assume that this justification also applies to the de Villiers' 1973 cross-sectional study of L-1 acquisition. However, to my knowledge, the use of the 90% criterion (or any other, for that matter) in L-2 research, whether longitudinal or cross-sectional, has not been justified. This questionable use of the 90% criterion for L-2 research is further complicated by the recent questionning of the validity of L-2 cross-sectional morpheme orders as acquisition orders. Larsen-Freeman (1975) and Bailey, Madden and Krashen (1974), for example, call the orders 'performance orders', 'accuracy orders', or 'difficulty orders'.

3. This was called the 90% Acquisition Index in Andersen (1976).

4. Andersen (1976) followed Dulay and Burt (1974a,b) in counting scores for those subjects who had three or more occasions to use the morpheme; in this paper scores are counted only if the individual had five or more occasions to use the morpheme. In Andersen (1976) the orders given by the 90% Acquisition Index and the Group Means Method correlated significantly at the .01 level, with a Pearson correlation coefficient of .98.

5. A word of caution is appropriate at this point. Since the subjects in this study had learned English mostly, or in many cases entirely, in compulsory ESL classes in a centralized public school system which requires uniform methods, textbooks and syllabuses in all public schools, it is very possible that the systematicity in the data revealed by the Group Range Method is due in part to this uniformity in the input. Thus, it is important that this study be replicated both within Puerto Rico and in other settings. Nevertheless, even if this systematicity should turn out to be a product of the highly homogeneous school system, the Group Range is still superior to the Group Means since it is sensitive enough to both the systematicity in group performance and variability within the group to pick up this uniformity.

6. COP and AUX mean ANY COP and ANY AUX. This is explained in greater detail on page 57.

7. I have chosen to use "+" in place of "1" (used by Bart and Krus and Dulay and Burt) so as to avoid confusion between percentages (1, 10, 11, etc.) and binary scores (+, + 0, + +, etc.).

8. These pairs of scores by no means exhaust the evidence in the data against binary scores. They are merely representative of a larger number of similar pairs of scores.

9. This is Bart and Krus' interpretation. (+ +) and (0 0) really neither confirm nor disconfirm the hypothesis.

10. This is not stated explicitly in either Bart and Krus (1973)or Dulay and Burt (1974a,b).

11. There were no cases of both raw percentage scores being 0%.

12. It remains to be seen whether the performance of the individual subjects on this specific task is an accurate reflection of their real "ability" or "knowledge" of the morphemes. Whether it is or not, binary scoring arbitrarily obscures real performance and should not be used without ample justification.

13. See Tarone, Frauenfelder and Selinker (1976:99-100) for a discussion of "learning strategies" vs."production strategies". It is not possible from the data in Table 7 to determine whether a learning strategy or a production strategy is involved in the reversing of AUX and ING.

14. I will not attempt to deal here with the need for a statistical test of significance of the percentage of subjects whose degree of fit is less dramatic but possibly significant. Such a test will be needed if we continue to use the test for individual fit.

15. Actually ANY AUX - ING, ANY AUX - V.

16. Treating bound morphemes and free morphemes separately may help explain morpheme orders (Krashen, personal communication).

17. This view of variability as also being systematic, which I have taken from Labov (1966 and in subsequent studies), is quite different from that given in Tarone, Frauenfelder and Selinker (1976:98) where variability is defined as being 'not systematic'. Tarone et al have given valuable insight into these notions of variability, systematicity and stability, but their view of variability needs considerable revision before it can fit the type of variability evident in this paper.

18. We might expect that a less experienced group of subjects would have more members in these cells.

19. Dulay and Burt (1974b:273-4 and in an erratum in Language Learning 25.1:207) briefly consider the relationship among these morphemes and two others, using the Ordering-Theoretic Method.

20. As in the preceding section, I am not dealing here with the question of whether a given percent of individual fit is statistically significant. This issue will eventually have to be dealt with.

21. For a less-experienced group of subjects we might expect to find more subjects in this last group.

22. I obviously have an advantage here working with written data. But even for my oral data it is usually possible to distinguish a from the.

23. Actually there are more contexts that require the definite article in Spanish than in English, as we shall see in 2.4. In addition, there are some contexts where Spanish requires the definite article and English requires possessive pronouns or something else.

24. Two recent longitudinal studies (Rosansky, 1976b:120-126; Hakuta, 1976: 338-341), which I had not read when I wrote this paper, come to basically the same conclusion that learners make more errors with A than with THE. (I am indebted to Steve Krashen for pointing this out to me and to Craig Chaudron for suggesting that I support the interpretation of Figure 3 with more data.) It is also worth noting that for Rosansky's subject, a Spanish speaker, THE attained criterion from the very beginning (tape 1), but A didn't attain criterion until tape 16 (approximately 8 months later), while for Hakuta's subject, a Japanese speaker, both THE and A attained criterion at the same point (sample 23), late in her acquisition of English.

Rosansky's longitudinal findings for one Spanish-speaking subject are similar to my cross-sectional findings for 89 Spanish-speaking subjects (see especially Rosansky's Figure 17 as compared with my Figure 3 and Hakuta's Figure 5). Since Spanish has articles but Japanese does not, Rosansky's conclusions seem to support my findings of L1 transfer in the use of articles. In fact she discusses this possibility in detail, especially with regard to A (p. 140-147) and concludes (147) that "this longitudinal study suggests that (for articles) interference is far more prevalant in the initial phases of L2 acquisition."

In Andersen (ms) I ordered the 89 subjects in my study in descending order of performance on all the morphemes I studied. I then divided this ordered series of subjects into 6 consecutive sub-samples A (subjects 1-15), B(16-30), C (31-45), D (46-60), E (61-75) and F (76-89). The Group Means for the entire sample (1-89) and for each sub-sample is given below for THE and A (from Andersen, ms).

		GROUP MEANS	
SUB-SAMPLE	SUBJECTS	THE	A
	1-89	96	84
A	1-15	98	92
B	16-30	98	92
C	31-45	95	85
D	46-60	98	85
E	61-75	95	75
F	76-89	92	71

It is evident from these figures that the less proficient the subjects are in the use of all the morphemes, the more errors they make in the use of A, whereas errors in using THE do not vary much and occur much less often. This adds support to Rosansky's tentative conclusions cited above as well as those of Taylor 1975 that L1 interference is more typical of the beginning stages of L2 acquisition. (Note, however, that Hakuta also reports that his subject "...performs better on THE than on A (340).")

My intention here, however, is not to argue for or against L1 transfer, but rather to try to clarify the interpretation of Figure 3 and to emphasize, once more, that the insight into the learning process provided by this study and Rosansky's and Hakuta's study of THE and A would not be possible if THE and A were combined as ARTICLE.

25. Actually errors of both inter- and intra-lingual interference as well as errors due to overcompensation (see 2.3).

APPENDIX Written Instructions Given to the Subjects in Spanish for Writing the Six Short Compositions

Tema 1 Comparación de dos casas

Compara dos casas que son muy diferentes la una de la otra. Por ejemplo, la casa de tus padres, (o de algún amigo o algún familiar) y la casa de tus abuelos. Compáralas por fuera y por dentro, incluyendo los diferentes objetos que se encuentran en cada una de las dos casas. Puedes comparar quizás una casa de campo con una casa en el pueblo, una casa sencílla con una más grande y más cara o cualesquier dos casas.

Tema 2 Descripción de acción

Escribe varios párrafos sobre lo que tú crees (o sabes o imaginas que varias personas diferentes están haciendo en este momento. Por ejemplo, tu mamá, tu papá, tus abuelos, tus tíos, tus amigos, tus hermanos, el gobernador de Puerto Rico, el presidente de Estados Unidos, alguna persona famosa, tus profesores, etc. Menciona por los menos 7 personas diferentes. Explica claramente quién es cada persona que mencionas (o sea, que es tu hermano, un amigo, un vecino, un tío, etc.).

Tema 3 Un fin de semana

Escribe varios párrafos sobre todo lo que hiciste el pasado fin de semana desde el viernes por la tarde hasta el domingo por la noche. Menciona cuantos detalles puedas, incluyendo una descripción de lo que hicieron o estaban haciendo otras personas que estaban contigo durante el fin de semana. Si no hiciste mucho el pasado fin de semana, escribe sobre algún otro fin de semana.

Topic 1 Comparison of two houses

Compare two houses that are very different from each other. For example, your parents' house (or the house of a friend or a member of your family) and your grandparents' house. Compare them on the outside and on the inside, including the different objects that are found in each of the two houses. You can compare, perhaps, a house in the country with a house in town, a simple house with a larger and more expensive house or any two houses.

Topic 2 Description of action

Write several paragraphs about what you think (or know or imagine) that several different persons are doing right now. For example, your mother, your father, your grandparents, your aunts and uncles, your friends, your brothers and sisters, the governor of Puerto Rico, the president of the United States, some famous person, your professors, etc. Mention at least 7 different people. Explain clearly who each person you mention is (that is, that he is your brother, a friend, a neighbor, an uncle, etc.).

Topic 3 A week-end

Write several paragraphs about everything you did last weekend from Friday afternoon until Sunday night. Mention as many details as you can, including a description of what other persons that were with you during the week-end did or were doing. If you didn't do much last week-end, write about any other week-end.

Tema 4 Un día típico en la vida de _____

Describe un día típico en la vida de alguien que tú conozcas muy bien. Por ejemplo, tu papá, tu mamá, tu abuelo, tu mejor amigo, etc. Escribe sobre las cosas que esa persona hace todos los días. Si no puedes escribir mucho sobre una de estas personas, escribe sobre dos o tres personas.

Topic 4 A typical day in the life of _____

Describe a typical day in the life of someone that you know very well. For example, your father, your mother, your grandfather, your best friend, etc. Write about the things that that person does every day. If you cannot write much about one of these persons, write about two or three persons.

Tema 5 Una entrevista con una persona famosa

En varios párrafos, escribe 10 ó 12 preguntas que le podrías hacer a alguna persona famosa. Escribe primero el nombre de la persona y luego escribe las preguntas.

1 3 ó 4 preguntas sobre el pasado de esa persona
2 3 ó 4 preguntas acerca de la vida actual de esa persona
3 3 ó 4 preguntas sobre los planes para el futuro de esa persona
4 3 ó 4 preguntas sobre diferentes cosas que esa persona ha hecho en su vida. Por ejemplo, le puedes preguntar si alguna vez ha visitado Puerto Rico o si ha estado muy infeliz alguna vez en su vida.

Topic 5 An interview with a famous person

In several paragraphs, write 10 or 12 questions that you could ask a famous person. Write first the name of the person and then write the questions.

1 3 or 4 questions about that person's past
2 3 or 4 questions about that person's present life
3 3 or 4 questions about that person's future plans
4 3 or 4 questions about different things that that person has done in his life. For example, you can ask him if he has ever visited Puerto Rico or if he has ever been unhappy anytime in his life.

Tema 6 Cosas que has hecho en tu vida

Escribe una oración completa en inglés para contestar cada una de estas preguntas.

1 ¿Has vivido alguna vez fuera de Puerto Rico?
2 ¿Has hecho alguna vez un viaje a Europa?
3 ¿Por cuántos años has estudiado el inglés?
4 ¿Alguna vez has amado a alguien que no te amó a ti?
5 ¿Has estado alguna vez en un accidente?
6 ¿Cuánto tiempo has estado en este salón hoy?
7 ¿Por cuántos años de tu vida has sido un estudiante?
8 ¿Has trabajado en algún lugar o siempre has sido estudiante?
9 ¿Has vivido toda tu vida en el mismo sitio?

Topic 6 Things you have done in your life

Write a complete sentence in English to answer each of these questions.

1 Have you ever lived outside of Puerto Rico?
2 Have you ever taken a trip to Europe?
3 For how many years have you studied English?
4 Have you ever loved someone that didn't love you?
5 Have you ever been in an accident?
6 How long have you been in this classroom today?
7 For how many years of your life have you been a student?
8 Have you ever worked some place or have you always been a student?
9 Have you lived in the same place all your life?

REFERENCES

Andersen, Roger W. 1976. A Function Acquisition Hierarchy Study in Puerto Rico. Paper presented at the 10th Annual TESOL Conference, New York: March, 1976.

_____ ms. A quantitative implicational model for second language acquisition.

Bailey, Nathalie, Carolyn Madden, and Stephen D. Krashen. 1974. Is there a "natural sequence" in adult second language learning? Language Learning 24.2:235-44.

Bart, William M. and David J. Krus. 1973. An Ordering-Theoretic Method to determine hierarchies among items. Educational and Psychological Measurement 33:291-300.

Brown, Roger. 1973. A First Language. Cambridge, Mass.: Harvard University Press.

De Villiers, Jill G. and Peter A. de Villiers. 1973. A cross-sectional study of the acquisition of grammatical morphemes in child speech. Journal of Psycholinguistic Research 2:267-78.

Dulay, Heidi and Marina Burt. 1974a. Natural sequences in child second language acquisition. Language Learning 24.1:37-53.

_____ 1974b. A new perspective on the creative construction process in child second language acquisition. Language Learning 24.2:287-97.

Hakuta, Kenji. 1974. Prefabricated patterns and the emergence of structure in second language acquisition. Language Learning 24.2:287-97.

_____ 1975. Learning to speak a second language: what exactly does the child learn? In: Georgetown University Round Table on Languages and Linguistics 1975. Edited by Daniel R. Dato. Washington, D.C.: Georgetown University Press. 195-207.

_____ 1976. A case study of a Japanese child learning English as a second language. Language Learning 26.2:321-51.

Krashen, Stephen D, Carolyn Madden and Nathalie Bailey. 1975. Theoretical aspects of grammatical sequencing. In: On TESOL 1975. Edited by Marina K. Burt and Heidi C. Dulay. 44-54.

Labov, William. 1966. The Social Stratification of English in New York City. Washington, D.C.: Center for Applied Linguistics.

Larsen-Freeman, Diane. 1975a. The acquisition of grammatical morphemes by adult ESL students. TESOL Quarterly Vol. 9, No. 4: 409-19.

_____ 1975b. The Acquisition of Grammatical Morphemes by Adult Learnings of English as a Second Language. Unpublished Ph.D. dissertation. The University of Michigan.

Rosansky, Ellen. 1976a. Morpheme studies and second language acquisition: a question of methods. Paper presented at the 10th Annual TESOL Conference, New York. March 1976.

_____ 1976b. Second Language Acquisition Research: A Question of Methods. Unpublished Ph.D. dissertation. Harvard.

Schachter, Jacquelyn. 1974. An error in error analysis. Language Learning 24.2:205-14.

Tarone, Elaine, Uli Frauenfelder, and Larry Selinker. 1976. Systematicity/Variability and Stability/Instability in Interlanguage Systems. In: Papers in Second Language Acquisition. Language Learning, Special Issue Num. 4. 93-134.

Taylor, Barry. 1975. The use of overgeneralization and transfer learning strategies by elementary and intermediate students in ESL. Language Learning 25.1:73-107.

Fathman, Ann K. 1975. Language Background, Age and the Order of Acquisition of English Structures. In On TESOL, '75: New Directions in Second Language Learning, Teaching and Bilingual Education. Edited by Marina K. Burt and Heidi C. Dulay. Washington, D.C.: TESOL.